

# 41



PCT09

ENTERED

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/462,972

DATE: 02/14/2003 <sup>86</sup>  
 TIME: 11:52:50

Input Set : A:\BB-1095-A Corrected Seq List.txt  
 Output Set: N:\CRF4\02142003\I462972.raw

3 <110> APPLICANT: E. I. DUPONT DE NEMOURS AND COMPANY  
 4 Odell, Joan T  
 5 Allen, Stephen M  
 7 \*120> TITLE OF INVENTION: PLANT SUG1 HOMOLOGS  
 9 <130> FILE REFERENCE: BB-1095-A  
 11 <140> CURRENT APPLICATION NUMBER: US 09/462,972  
 12 <141> CURRENT FILING DATE: 2000-01-14  
 14 <150> PRIOR APPLICATION NUMBER: PCT/US98/13992  
 15 <151> PRIOR FILING DATE: 1998-07-07  
 17 <150> PRIOR APPLICATION NUMBER: US 08/893,401  
 18 <151> PRIOR FILING DATE: 1997-07-11  
 20 <160> NUMBER OF SEQ ID NOS: 14  
 22 <170> SOFTWARE: Microsoft Office 97  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 1254  
 26 <212> TYPE: DNA  
 27 <213> ORGANISM: Glycine max  
 29 <220> FEATURE:  
 30 <221> NAME/KEY: CDS  
 31 <222> LOCATION: (1)..(1254)  
 32 <223> OTHER INFORMATION:  
 W--> 34 <400> 1  
 35 atg gct ctt gta gga gtt gaa ctg aag cat gcg gcg gag ggc gta ccg 48  
 36 Met Ala Leu Val Gly Val Glu Leu Lys His Ala Ala Glu Gly Val Pro 15  
 37 1 5 96  
 39 gag gcg aat tgc tcc gcc aag ccc acc aag cag ggc gag ggc ctc cgc  
 40 Glu Ala Asn Cys Ser Ala Lys Pro Thr Lys Gln Gly Glu Gly Leu Arg 30  
 41 20 144  
 43 cac tac tat tct ctc aac atc cac gag cat cag ctc ctt ctt cgc caa  
 44 His Tyr Tyr Ser Leu Asn Ile His Glu His Gln Leu Leu Arg Gln 45  
 45 35 192  
 47 aag act cat aac ctc aac cgt ctc gag gct cag aga aac gac ctc aat  
 48 Lys Thr His Asn Leu Asn Arg Leu Glu Ala Gln Arg Asn Asp Leu Asn 60  
 49 50 240  
 51 tct agg gtg agg atg ctg cgc gaa gaa tta cag ctt ctg cag gaa ccc  
 52 Ser Arg Val Arg Met Leu Arg Glu Glu Leu Gln Leu Leu Gln Glu Pro 80  
 53 65 70 75 288  
 55 ggc tct tat gtc ggt gaa gtt gtc aaa gta atg ggc aag aac aaa gtc  
 56 Gly Ser Tyr Val Gly Glu Val Val Lys Val Met Gly Lys Asn Lys Val 95  
 57 85 90 336  
 59 ctt gtc aag gtc cac cca gaa gga aaa tat gtt gtt gac att gac aaa  
 60 Leu Val Lys Val His Pro Glu Gly Lys Tyr Val Val Asp Ile Asp Lys 110  
 61 100

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/462,972

DATE: 02/14/2003  
TIME: 11:52:50Input Set : A:\BB-1095-A Corrected Seq List.txt  
Output Set: N:\CRF4\02142003\I462972.raw

63	aat att gac att aca aag att act cca tcc act aga gtt gca ctc cgc	384
64	Asn Ile Asp Ile Thr Lys Ile Thr Pro Ser Thr Arg Val Ala Leu Arg	
65	115 120	
67	aac gac agt tat gtt ctt cac tta gtt ctg cca agt aaa gtt gat cca	432
68	Asn Asp Ser Tyr Val Leu His Leu Val Leu Pro Ser Lys Val Asp Pro	
69	130 135	
71	ttg gtc aat ctg atg aaa gtt gag aaa gtt ccc gat tct aca tat gac	480
72	Leu Val Asn Leu Met Lys Val Glu Lys Val Pro Asp Ser Thr Tyr Asp	
73	145 150	
75	atg att ggt ggt tta gac cag caa att aaa gaa ata aaa gag gtc att	528
76	Met Ile Gly Gly Leu Asp Gln Gln Ile Lys Glu Ile Lys Glu Val Ile	
77	165 170	
79	gag cta cca atc aaa cat cct gag ctg ttt gaa agt ctt gga att gca	576
80	Glu Leu Pro Ile Lys His Pro Glu Leu Phe Glu Ser Leu Gly Ile Ala	
81	180 185	
83	caa cca aag ggt gtc ctg ctc tat ggg cca cct ggt aca ggt aaa aca	624
84	Gln Pro Lys Gly Val Leu Leu Tyr Gly Pro Pro Gly Thr Gly Lys Thr	
85	195 200	
87	ttg ttg gct agg gca gtg gct cat cat act gac tgt aca ttc atc agg	672
88	Leu Leu Ala Arg Ala Val His His Thr Asp Cys Thr Phe Ile Arg	
89	210 215	
91	gtg tct ggt tct gag tta gtt cag aaa tac att gga gaa ggt tct aga	720
92	Val Ser Gly Ser Glu Leu Val Gln Lys Tyr Ile Gly Glu Gly Ser Arg	
93	225 230	
95	atg gtc agg gaa ctt ttt gtt atg gcc agg gaa cat gct cca tca att	768
96	Met Val Arg Glu Leu Phe Val Met Ala Arg Glu His Ala Pro Ser Ile	
97	245 250	
99	atc ttc atg gat gaa att gac agt att gga tct gct cgg atg gaa tct	816
100	Ile Phe Met Asp Glu Ile Asp Ser Ile Gly Ser Ala Arg Met Glu Ser	
101	260 265	
103	gga agt ggc aac ggt gat agt gag gta cag cgt act atg ctg gaa ctt	864
104	Gly Ser Gly Asn Gly Asp Ser Glu Val Gln Arg Thr Met Leu Glu Leu	
105	275 280	
107	ctc aac cag ttg gat gga ttt gaa gct tca aat aag atc aag gtt ttg	912
108	Leu Asn Gln Leu Asp Gly Phe Glu Ala Ser Asn Lys Ile Lys Val Leu	
109	290 295	
111	atg gcc acc aat cgg att gat atc ctg gat caa gcc ctc ctt aga cca	960
112	Met Ala Thr Asn Arg Ile Asp Ile Leu Asp Gln Ala Leu Leu Arg Pro	
113	305 310	
115	gga cgg ata gac cgg aaa att gaa ttt cca acc cct aat gaa gag tct	1008
116	Gly Arg Ile Asp Arg Lys Ile Glu Phe Pro Thr Pro Asn Glu Glu Ser	
117	325 330	
119	cgg ctg gat att ttg aaa atc cat tct aga aga atg aat tta atg cgt	1056
120	Arg Leu Asp Ile Leu Lys Ile His Ser Arg Arg Met Asn Leu Met Arg	
121	340 345	
123	ggc att gat ttg aag aag att gcc gag aag atg aat gga gca tct ggt	1104
124	Gly Ile Asp Leu Lys Lys Ile Ala Glu Lys Met Asn Gly Ala Ser Gly	
125	355 360	
127	gct gaa ctt aag gct gtt tgc act gaa gct gga atg ttt gct ttg agg	1152

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/462,972

DATE: 02/14/2003  
TIME: 11:52:50Input Set : A:\BB-1095-A Corrected Seq List.txt  
Output Set: N:\CRF4\02142003\I462972.raw

```

128 Ala Glu Leu Lys Ala Val Cys Thr Glu Ala Gly Met Phe Ala Leu Arg
129      370      375      380      1200
131 gag cgg agg gta cac gtg act cag gag gat ttt gag atg gcc gtg gcg
132 Glu Arg Arg Val His Val Thr Gln Glu Asp Phe Glu Met Ala Val Ala
133 385      390      395      400      1248
135 aag gtg atg aaa aag gag act gaa aaa aac atg tca ttg cgg aag ttg
136 Lys Val Met Lys Lys Glu Thr Glu Lys Asn Met Ser Leu Arg Lys Leu
137      405      410      415      1254
139 tgg aag
140 Trp Lys
144 <210> SEQ ID NO: 2
145 <211> LENGTH: 418
146 <212> TYPE: PRT
147 <213> ORGANISM: Glycine max
149 <400> SEQUENCE: 2
151 Met Ala Leu Val Gly Val Glu Leu Lys His Ala Ala Glu Gly Val Pro
152 1      5      10      15
155 Glu Ala Asn Cys Ser Ala Lys Pro Thr Lys Gln Gly Glu Leu Arg
156      20      25      30
159 His Tyr Tyr Ser Leu Asn Ile His Glu His Gln Leu Leu Arg Gln
160      35      40      45
163 Lys Thr His Asn Leu Asn Arg Leu Glu Ala Gln Arg Asn Asp Leu Asn
164      50      55      60
167 Ser Arg Val Arg Met Leu Arg Glu Glu Leu Gln Leu Leu Gln Glu Pro
168 65      70      75      80
171 Gly Ser Tyr Val Gly Glu Val Val Lys Val Met Gly Lys Asn Lys Val
172      85      90      95
175 Leu Val Lys Val His Pro Glu Gly Lys Tyr Val Val Asp Ile Asp Lys
176      100      105      110
179 Asn Ile Asp Ile Thr Lys Ile Thr Pro Ser Thr Arg Val Ala Leu Arg
180      115      120      125
183 Asn Asp Ser Tyr Val Leu His Leu Val Leu Pro Ser Lys Val Asp Pro
184      130      135      140
187 Leu Val Asn Leu Met Lys Val Glu Lys Val Pro Asp Ser Thr Tyr Asp
188 145      150      155      160
191 Met Ile Gly Gly Leu Asp Gln Gln Ile Lys Glu Ile Lys Glu Val Ile
192      165      170      175
195 Glu Leu Pro Ile Lys His Pro Glu Leu Phe Glu Ser Leu Gly Ile Ala
196      180      185      190
199 Gln Pro Lys Gly Val Leu Leu Tyr Gly Pro Pro Gly Thr Gly Lys Thr
200      195      200      205
203 Leu Leu Ala Arg Ala Val Ala His His Thr Asp Cys Thr Phe Ile Arg
204      210      215      220
207 Val Ser Gly Ser Glu Leu Val Gln Lys Tyr Ile Gly Glu Gly Ser Arg
208 225      230      235      240
211 Met Val Arg Glu Leu Phe Val Met Ala Arg Glu His Ala Pro Ser Ile
212      245      250      255
215 Ile Phe Met Asp Glu Ile Asp Ser Ile Gly Ser Ala Arg Met Glu Ser
216      260      265      270

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/462,972

DATE: 02/14/2003  
TIME: 11:52:50Input Set : A:\BB-1095-A Corrected Seq List.txt  
Output Set: N:\CRF4\02142003\I462972.raw

219 Gly Ser Gly Asn Gly Asp Ser Glu Val Gln Arg Thr Met Leu Glu Leu  
220 275 280  
223 Leu Asn Gln Leu Asp Gly Phe Glu Ala Ser Asn Lys Ile Lys Val Leu  
224 290 295  
227 Met Ala Thr Asn Arg Ile Asp Ile Leu Asp Gln Ala Leu Leu Arg Pro  
228 305 310  
231 Gly Arg Ile Asp Arg Lys Ile Glu Phe Pro Thr Pro Asn Glu Glu Ser  
232 325 330  
235 Arg Leu Asp Ile Leu Lys Ile His Ser Arg Arg Met Asn Leu Met Arg  
236 340 345  
239 Gly Ile Asp Leu Lys Lys Ile Ala Glu Lys Met Asn Gly Ala Ser Gly  
240 355 360  
243 Ala Glu Leu Lys Ala Val Cys Thr Glu Ala Gly Met Phe Ala Leu Arg  
244 370 375  
247 Glu Arg Arg Val His Val Thr Gln Glu Asp Phe Glu Met Ala Val Ala  
248 385 390  
251 Lys Val Met Lys Lys Glu Thr Glu Lys Asn Met Ser Leu Arg Lys Leu  
252 405 410  
255 Trp Lys

259 &lt;210&gt; SEQ ID NO: 3

260 &lt;211&gt; LENGTH: 1148

261 &lt;212&gt; TYPE: DNA

262 &lt;213&gt; ORGANISM: Zea mays

264 &lt;220&gt; FEATURE:

265 &lt;221&gt; NAME/KEY: CDS

266 &lt;222&gt; LOCATION: (3)..(1148)

267 &lt;223&gt; OTHER INFORMATION:

W--&gt; 269 &lt;400&gt; 3

270 ga gag cac atc cat gac ctg cag ctc cag atc cgg cag aag acc cat  
271 Glu His Ile His Asp Leu Gln Leu Gln Ile Arg Gln Lys Thr His  
272 1 5 10  
274 aac ctc aac cgc ctc gag gcc cag cgc aac gac ctc aac tcc cga gtt  
275 Asn Leu Asn Arg Leu Glu Ala Gln Arg Asn Asp Leu Asn Ser Arg Val  
276 20 25 30  
278 aga atg ctc agg gaa gag ttg cag ttg ctt caa gag cct ggc tca tat  
279 Arg Met Leu Arg Glu Glu Leu Gln Leu Leu Gln Glu Pro Gly Ser Tyr  
280 35 40 45  
282 gtt ggt gag gtg gtg aag gtc atg ggg aaa tca aag gtt ctg gtg aag  
283 Val Gly Glu Val Val Lys Val Met Gly Lys Ser Lys Val Leu Val Lys  
284 50 55 60  
286 gta cat ccc gaa ggc aaa tat gtg gtg gat ata gat aag agc att gat  
287 Val His Pro Glu Gly Lys Tyr Val Val Asp Ile Asp Lys Ser Ile Asp  
288 65 70 75  
290 atc act aag atc aca cct tca aca aga gtt gct ctt cgg aat gac agc  
291 Ile Thr Lys Ile Thr Pro Ser Thr Arg Val Ala Leu Arg Asn Asp Ser  
292 80 85 90  
294 tat atg ctc cat ctg atc cta cca agc aaa gtt gat cca ttg gtc aat  
295 Tyr Met Leu His Leu Ile Leu Pro Ser Lys Val Asp Pro Leu Val Asn  
296 100 105 110

47

95

143

191

239

287

335

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/462,972

DATE: 02/14/2003  
TIME: 11:52:50

Input Set : A:\BB-1095-A Corrected Seq List.txt  
Output Set: N:\CRF4\02142003\I462972.raw

298	ctc atg aaa gtt gag aag gtt ccg gat tct acc tat gat atg att gga	383
299	Leu Met Lys Val Glu Lys Val Pro Asp Ser Thr Tyr Asp Met Ile Gly	
300	115 120 125	
302	ggc ctt gac cag caa att aaa gag atc aaa gag gtc att gag ctt cca	431
303	Gly Leu Asp Gln Gln Ile Lys Glu Ile Lys Glu Val Ile Glu Leu Pro	
304	130 135 140	
306	atc aaa cat ccg gaa ctg ttt gag agc ctt gga att gcg caa cca aag	479
307	Ile Lys His Pro Glu Leu Phe Glu Ser Leu Gly Ile Ala Gln Pro Lys	
308	145 150 155	
310	ggt gtc ctt ctt tat gga cct ccg ggc aca gga aag aca ttg ttg gca	527
311	Gly Val Leu Leu Tyr Gly Pro Pro Gly Thr Gly Lys Thr Leu Leu Ala	
312	160 165 170 175	
314	cgt gcg gtt gct cat cac act gac tgc acc ttc atc agg gtg tct ggt	575
315	Arg Ala Val Ala His His Thr Asp Cys Thr Phe Ile Arg Val Ser Gly	
316	180 185 190	
318	tct gag ttg gtt cag aag tat att ggt gag ggc tcc cgg atg gtt agg	623
319	Ser Glu Leu Val Gln Lys Tyr Ile Gly Glu Gly Ser Arg Met Val Arg	
320	195 200 205	
322	gaa ctc ttt gtt atg gcc agg gaa cat gca cca tcc att ata ttt atg	671
323	Glu Leu Phe Val Met Ala Arg Glu His Ala Pro Ser Ile Ile Phe Met	
324	210 215 220	
326	gat gaa att gac tct atc gga tct gct aga atg gag tct gga act ggc	719
327	Asp Glu Ile Asp Ser Ile Gly Ser Ala Arg Met Glu Ser Gly Thr Gly	
328	225 230 235	
330	aac ggt gac agt gaa gtt cag cgt act atg ctt gaa ctt cta aac cag	767
331	Asn Gly Asp Ser Glu Val Gln Arg Thr Met Leu Glu Leu Leu Asn Gln	
332	240 245 250 255	
334	ctc gat ggt ttt gaa gca tca aac aaa att aag gtt ttg atg gca acg	815
335	Leu Asp Gly Phe Glu Ala Ser Asn Lys Ile Lys Val Leu Met Ala Thr	
336	260 265 270	
338	aac aga ata gac att ttg gat caa gcc ctt ctg agg cct ggc cgc ata	863
339	Asn Arg Ile Asp Ile Leu Asp Gln Ala Leu Leu Arg Pro Gly Arg Ile	
340	275 280 285	
342	gac agg aag att gaa ttt cca aat cct aac gag gat tca cgt ttc gat	911
343	Asp Arg Lys Ile Glu Phe Pro Asn Pro Asn Glu Asp Ser Arg Phe Asp	
344	290 295 300	
346	atc ttg aag atc cat tca aga aaa atg aac ttg atg cgt ggc att gat	959
347	Ile Leu Lys Ile His Ser Arg Lys Met Asn Leu Met Arg Gly Ile Asp	
348	305 310 315	
350	ctg aaa aag atc gcg gaa aag atg aat ggg gcc tca gga gct gag ctc	1007
351	Leu Lys Lys Ile Ala Glu Lys Met Asn Gly Ala Ser Gly Ala Glu Leu	
352	320 325 330 335	
354	aag gcc gtc tgc aca gag gct gga atg ttt gct ctt cgt gag aga agg	1055
355	Lys Ala Val Cys Thr Glu Ala Gly Met Phe Ala Leu Arg Glu Arg Arg	
356	340 345 350	
358	gtg cac gtt acc cag gag gac ttc gag atg gca gtg gcc aag gtg atg	1103
359	Val His Val Thr Gln Glu Asp Phe Glu Met Ala Val Ala Lys Val Met	
360	355 360 365	
362	aag aaa gac acg gag aag aac atg tcc ctg cgc aag ctc tgg aag	1148

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/462,972

DATE: 02/14/2003  
TIME: 11:52:51

Input Set : A:\BB-1095-A Corrected Seq List.txt  
Output Set: N:\CRF4\02142003\I462972.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:11; N Pos. 262,271,293,331,355,356,373,384,387,391,398,405,407,418,425  
Seq#:11; N Pos. 427,428,453,454,473,497,518,526,527,541  
Seq#:12; Xaa Pos. 78,79,82,90,110,111  
Seq#:13; N Pos. 149,206,217,299,325,354,412,478,493,498,521,558,571,584,585  
Seq#:14; Xaa Pos. 20,47,66,70,97,101,106,115,116

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/462,972

DATE: 02/14/2003

TIME: 11:52:51

Input Set : A:\BB-1095-A Corrected Seq List.txt  
Output Set: N:\CRF4\02142003\I462972.raw

L:34 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:32  
L:269 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:267  
L:1222 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:240  
M:341 Repeated in SeqNo=11  
L:1283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:64  
M:341 Repeated in SeqNo=12  
L:1393 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:120  
M:341 Repeated in SeqNo=13  
L:1469 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:16  
M:341 Repeated in SeqNo=14